

XX	10	18	50.6	35099	19	AAV27112	Adenovirus 17. Me
XX	11	18.4	59.4	415	21	AAO29348	Fusarium venenatum
XX	12	18.4	59.4	490	22	AAH12307	Human cDNA clone (
XX	13	18.4	59.4	3894	23	ABL19721	Drosophila melanog
XX	14	18.4	59.4	4263	23	ABL15567	Drosophila melanog
XX	15	18.4	59.4	6028	23	ABL19720	Drosophila melanog
XX	16	18.4	59.4	13499	23	ABL15566	Drosophila melanog
XX	17	18.2	58.7	60	18	AAI49396	Banana bunchy top
XX	18	18.2	58.7	69	18	AAI49393	Banana bunchy top
XX	19	18.2	58.7	206	22	ABA47731	Human breast cell
XX	20	18.2	58.7	206	22	ABA47731	Human breast cell
XX	21	18.2	58.7	206	22	ABA65625	Human foetal liver
XX	22	18.2	58.7	206	22	ABA65625	Human foetal liver
XX	23	18.2	58.7	206	22	ABA32712	Probe #11178 for g
XX	24	18.2	58.7	206	22	ABA32712	Probe #11178 for g
XX	25	18.2	58.7	206	22	AAK14028	Human brain expres
XX	26	18.2	58.7	206	22	AAK14028	Human brain expres
XX	27	18.2	58.7	206	22	AAK39774	Human bone marrow
XX	28	18.2	58.7	206	22	AAK39774	Human bone marrow
XX	29	18.2	58.7	206	22	AAI20582	Probe #10515 for g
XX	30	18.2	58.7	206	22	AAI20582	Probe #10515 for g
XX	31	18.2	58.7	206	22	AAI45791	Probe #14477 used
XX	32	18.2	58.7	206	22	AAI45791	Probe #14477 used
XX	33	18.2	58.7	206	22	AAI06281	Probe #6272 used t
XX	34	18.2	58.7	206	22	AAI06281	Probe #6272 used t
XX	35	18.2	58.7	352	18	ABA49407	Banana bunchy top
XX	36	18.2	58.7	478	22	ABA42602	Human breast cell
XX	37	18.2	58.7	478	22	ABA42602	Human breast cell
XX	38	18.2	58.7	478	22	ABA53031	Human foetal liver
XX	39	18.2	58.7	478	22	ABA53031	Human foetal liver
XX	40	18.2	58.7	478	22	ABA22805	Probe #1271 for ge
XX	41	18.2	58.7	478	22	ABA22805	Probe #1271 for ge
XX	42	18.2	58.7	478	22	AAK01279	Human brain expres
XX	43	18.2	58.7	478	22	AAK01279	Human brain expres
XX	44	18.2	58.7	478	22	AAK26739	Human bone marrow
XX	45	18.2	58.7	478	22	AAK26739	Human bone marrow

XX	WPI; 1994-058406/08.
DR	Use of coconut foliar decay virus DNA as promoter - for
XX	tissue-specific gene expression in transgenic plants
PT	Claim 1; Fig 2; 8pp; German.
XX	A DNA fragment from the CPDV genome can be used as a phloem-specific
CC	promoter in the construction of transgenic plants. The promoter is
CC	strong; it has 30-50% of the activity of the CamV 35S promoter in
CC	tabacco plants.
XX	Sequence 31 BP; 5 A; 11 C; 11 G; 4 T; 0 other;
SQ	
Query Match	100.0%; Score 31; DB 15; Length 31;
Best Local Similarity	100.0%; Pred. No. 9,5e-05;
Matches	31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 agccgcggggtgaataactagcccccgagct 31
Dd	1 agccgcggggtgaataactagcccccgagct 31
RESULT 2	
AAx02688	
ID	AAx02688 standard; DNA; 31 BP.
XX	AAx02688;
AC	
XX	10-MAY-1999 (first entry)
DI	
XX	Gemini virus DNA fragment stem loop.
DE	
XX	Stem loop; coconut foliar decay virus; CPDV; bacterial promoter;
KM	yeast promoter; tissue-specific; phloem-specific; CamV 35S promoter; ss.
XX	
OS	Gemini virus.
XX	
FH	Key Location/Qualifiers
FT	stem_loop 1..31
FT	/tag= a
FT	misc_binding 1..11
FT	/tag= b
FT	/note= "Region binds to nucleotides 21 to 31"
FT	misc_binding 21..31
FT	/tag= c
FT	/note= "Region binds to nucleotides 1 to 11"
PN	
XX	DE19730502-A1.
XX	
PD	21-JAN-1999.
XX	
PF	16-JUL-1997; 97DE-1030502.
XX	
PR	16-JUL-1997; 97DE-1030502.
XX	
PA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX	
PI	Becker D, Hehn A, Randles JW, Rohde W, Salamini F;
XX	WPI; 1999-096863/09.
DR	
XX	Cocoon foliar decay virus promoters - for gene expression in
PT	bacteria and yeasts
XX	
PS	Disclosure; Fig 2; 14pp; German.
XX	
CC	This invention describes a coconut foliar decay virus (CPDV) DNA fragment
CC	that includes the stem-loop structure of CPDV DNA but lacks the
CC	translation start codons of open reading frames ORF1 and/or ORF2. The
CC	new CPDV DNA fragment is useful as a bacterial or yeast promoter, as a

```
CC promoter for tissue-specific (especially phloem-specific) gene expression
CC in plants and for production of chimeric constructs for transient or
CC stable expression. Certain fragments of CPDV DNA have stronger promoter
CC activity in E. coli than the CamV 35S promoter.
XX
SQ Sequence 31 BP; 5 A; 11 C; 11 G; 4 T; 0 other;

Query Match          100.0%; Score 31; DB 20; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 agccgcggggaatcaactagccccgcgcgt 31
        |||
Db       1 agccgcggggaatactagccccgcgcgt 31

RESULT      3
AAK02686
ID   AAK02686 standard; DNA; 31 BP.
AC   AAK02686;
XX
XX      10-MAY-1999 (first entry)
DT
DE      Gemini virus DNA fragment stem loop.
XX
XX      Stem loop: coconut foliar decay virus; CPDV; bacterial promoter;
KW      yeast promoter; tissue-specific; phloem-specific; CamV 35S promoter; ss.
OS      Gemini virus.
XX
FH      Key Location/Qualifiers
FT      stem_loop            1..31
FT                                 /*tag= a
FT      misc_binding         1..11
FT                                 /*tag= b
FT      misc_binding         /note= "Region binds to nucleotides 21 to 31"
FT      misc_binding         21..31
FT                                 /*tag= c
FT                                 /note= "Region binds to nucleotides 1 to 11"
FT
PN      DE19730535-A1.
XX
XX      21-JAN-1999.
PD
PF      16-JUL-1997;    97DE-1030535.
XX
PR      16-JUL-1997;    97DE-1030535.
XX
PA      (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX      Becker D, Hehn A, Randles JW, Rohde W, Salamini F;
PI      WPI; 1999-096867/09.
DR
XX      Coconut foliar decay virus promoters - for gene expression in
PT      plants, bacteria and yeasts
XX
PS      Disclosure; Fig 2; 12pp; German.
XX
XX      This invention describes a coconut foliar decay virus (CPDV) DNA fragment
XX      that includes the stem-loop structure of CPDV DNA but lacks the
XX      translation start codons of open reading frames ORF1 and/or ORF2. The
XX      new CPDV DNA fragment is useful as a bacterial or yeast promoter, as a
XX      promoter for tissue-specific (especially phloem-specific) gene expression
XX      in plants and for production of chimeric constructs for transient or
XX      stable expression. Certain fragments of CPDV DNA have stronger promoter
XX      activity in E. coli than the CamV 35S promoter.
XX
SQ Sequence 31 BP; 5 A; 11 C; 11 G; 4 T; 0 other;
```


PR	01-DEC-2000;	2000US-0250391.
PR	01-DEC-2000;	2000US-0251160.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	11-DEC-2000;	2000US-0251990.
PR	05-JAN-2001;	2001US-0254097.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
DR	WPI; 2001-541565/60.	
XX		
PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating nervous system	
PT	cancers and metastases -	
PS		
PX	Disclousure; SEQ ID NO 8740; 1701pp + Sequence Listing; English.	
CC	The invention relates to novel genes (ABA11004-ABR21534) and proteins	
CC	(ABR14678-ABR18001) useful for preventing, treating or ameliorating	
CC	medical conditions e.g. by protein or gene therapy. The genes are	
CC	isolated from a range of human tissues disclosed in the specification.	
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful	
CC	in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast	
CC	and ovarian cancer and other cancers of the adrenal gland, bone, bone	
CC	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;	
CC	(b) immune disorders e.g. Addison's disease, allergies, autoimmune	
CC	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's	
CC	disease, multiple sclerosis, rheumatoid arthritis and ulcerative	
CC	colitis; (c) cardiovascular disorders such as myocardial ischaemias;	
CC	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and	
CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal	
CC	and parasitic infections.	
CC	Note: The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 1423 BP; 253 A; 397 C; 425 G; 348 T; 0 other;	
Query Match	66.5%; Score 20.6; DB 22; Length 1423;	
Best Local Similarity	85.2%; Pred. No. 6.7;	
Matches 23; Conservative	0; Mismatches 4; Indels 0; Gaps 0;	
OY	1 agccgcggggtactactagccccgc 27	
DG		
DB	141 AGCCGCCGGGGTGAGTCAGCCCCAGC 115	
RESULT	6	
AAO56057/c		
ID	AAO56057 standard; DNA; 31 BP.	
XX		
AC	AAO56057;	
XX		
DT	12-AUG-1994 (first entry)	
XX		
DE	Coconut Foliar Decay Virus promoter stem-loop.	
XX		
KM	Coconut Foliar Decay Virus; CFDV; strong promoter; tissue-specific;	
KW	phloem-specific; stem-loop structure; transgenic plant; ds.	
XX		
OS	Coconut Foliar Decay Virus.	
XX		
Key	Location/Qualifiers	

```

FT      stem_loop      1..31
FT      /*tag= a
FT      /function= promoter
FT      /note= "loop has homology to geminivirus sequence"
XX
XX      DE4306832-C.
XX
XX      24-FEB-1994.
XX
XX      04-MAR-1993; 93DE-4306832.
XX
XX      04-MAR-1993; 93DE-4306832.
XX
XX      (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX      Becker D, Randles JW, Rohde W, Salamini F;
XX      WPI; 1994-058406/08.
XX
XX      Use of coconut foliar decay virus DNA as promoter - for
XX      tissue-specific gene expression in transgenic plants
XX
XX      Claim 1; Fig 2; 8pp; German.
XX
XX      A DNA fragment from the CPDV genome can be used as a phloem-specific
XX      promoter in the construction of transgenic plants. The promoter is
XX      strong; it has 30-50% of the activity of the CaMV 35S promoter in
XX      tobacco plants.
XX
XX      Sequence 31 BP; 5 A; 11 C; 11 G; 4 T; 0 other;
XX

```

```

Query Match      63.9%; Score 19.8; DB 15; Length 31;
Best Local Similarity 77.4%; Pred. No. 11;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

OY      1 agccgcgggggtactactagccccgcgct 31
      |||||
      31 AGCCGCGGGGCTAGTATTACCCCGCGCT 1

```

```

RESULT 7
AAAX02688/C
AAAX02688 standard; DNA: 31 BP.

```

```

AC      AAX02688;

```

```

XX      10-MAY-1999 (first entry)

```

```

DE      Gemini virus DNA fragment stem loop.

```

```

XX      Stem loop; coconut foliar decay virus; CPDV; bacterial promoter;
KW      yeast promoter; tissue-specific; phloem-specific; CaMV 35S promoter; ss.
XX

```

```

OS      Gemini virus.

```

```

XX      Location/Qualifiers

```

```

FT      Key      1..31
FT      stem_loop      /*tag= a
FT      misc_binding      1..11
FT      /*tag= b
FT      /note= "Region binds to nucleotides 21 to 31"
FT      misc_binding      21..31
FT      /*tag= c
FT      /note= "Region binds to nucleotides 1 to 11"

```

```

XX      DE19730502-A1.

```

```

XX      21-JAN-1999.

```

```

XX      16-JUL-1997; 97DE-1030502.

```

```

XX      16-JUL-1997; 97DE-1030502.

```

```

XX      (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX      Becker D, Hehn A, Randles JW, Rohde W, Salamini F;
XX      WPI; 1999-096863/09.
XX
XX      Coconut foliar decay virus promoters - for gene expression in
XX      bacteria and yeasts
XX
XX      Disclosure; Fig 2; 14pp; German.
XX
XX      This invention describes a coconut foliar decay virus (CPDV) DNA fragment
XX      that includes the stem-loop structure of CPDV DNA but lacks the
XX      translation start codons of open reading frames ORF1 and/or ORF2. The
XX      new CPDV DNA fragment is useful as a bacterial or yeast promoter, as a
XX      promoter for tissue-specific (especially phloem-specific) gene expression
XX      in plants and for production of chimeric constructs for transient or
XX      stable expression. Certain fragments of CPDV DNA have stronger promoter
XX      activity in E. coli than the CaMV 35S promoter.
XX
XX      Sequence 31 BP; 5 A; 11 C; 11 G; 4 T; 0 other;
XX

```

```

Query Match      63.9%; Score 19.8; DB 20; Length 31;
Best Local Similarity 77.4%; Pred. No. 11;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

OY      1 agccgcgggggtactactagccccgcgct 31
      |||||
      31 AGCCGCGGGGCTAGTATTACCCCGCGCT 1

```

```

RESULT 8
AAAX02686/C
AAAX02686 standard; DNA: 31 BP.

```

```

AC      AAX02686;

```

```

XX      10-MAY-1999 (first entry)

```

```

DE      Gemini virus DNA fragment stem loop.

```

```

XX      Stem loop; coconut foliar decay virus; CPDV; bacterial promoter;
KW      yeast promoter; tissue-specific; phloem-specific; CaMV 35S promoter; ss.
XX

```

```

OS      Gemini virus.

```

```

XX      Location/Qualifiers

```

```

FT      Key      1..31
FT      stem_loop      /*tag= a
FT      misc_binding      1..11
FT      /*tag= b
FT      /note= "Region binds to nucleotides 21 to 31"
FT      misc_binding      21..31
FT      /*tag= c
FT      /note= "Region binds to nucleotides 1 to 11"
XX
XX      DE19730535-A1.
XX
XX      21-JAN-1999.
XX
XX      16-JUL-1997; 97DE-1030535.
XX
XX      16-JUL-1997; 97DE-1030535.
XX
XX      16-JUL-1997; 97DE-1030535.
XX

```

```

XX      (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

```

```

XX      Becker D, Hehn A, Randles JW, Rohde W, Salamini F;

```

```

XX      WPI; 1999-096867/09.

```

```

XX      Coconut foliar decay virus promoters - for gene expression in

```

CC a food supplement. (II) and its binding partners are useful in medical

CC penton bases such as Ad17 penton base (see AAV27114). The chimeric
CC vector is used to target biologically active proteins to airway
CC epithelial cells, especially for gene therapy of diseases such as
CC cystic fibrosis or alpha-antitrypsin deficiency. The vectors are


```

XX ABL19720;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 10633.
DE
XX Drosophila: developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
PA
XX (PEKE ) PE CORP NY.
PI
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PS
XX Claim 1; SEQ ID NO 10633; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL1840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
CC
XX
XX Sequence 6028 BP; 1486 A; 1441 C; 1361 G; 1740 T; 0 other;
SQ

```

```

Query Match          59.48; Score 18.4; DB 23; Length 6028;
Best Local Similarity 78.68; Pred. No. 75;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 2 gccgcgggggtatatactagcccccgcgg 29
   ||||| ||||| ||||| ||||| |||||
DB 3883 GCGGCGGAGGAAAACTAGCCGCCGAGG 3856

```

```

Search completed: June 25, 2002, 14:42:05
Job time: 146 sec

```

```
source
1. .905
/organism="Tetraodon nigroviridis"
```



```

source
1. .1843
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:403511"
/clone_1ib="NIH-MGC_46"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: XhoI; EcoRI: CDNA made by oligo-dT priming into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."
BASE COUNT      262 a      561 c      500 g      516 t      4 others
ORIGIN

Query Match      64.5%      Score 20;      DB 10;      Length 1843;
Best Local Similarity 82.1%;      Pred. No. 6.6e+02;
Matches 23;      Conservative 0;      Mismatches 5;      Indels 0;      Gaps 0;

Qy      4      cgcggggaatcactacgccccgcgcgt 31
      ||||| |||| | |||| |||||
Db      1609      CGCGGGGTGAAGAAGTAGCTCCACCGGCT 1582

RESULT      6
BE902577      865 bp      mRNA      linear      EST 29-SEP-2000
LOCUS      60167772571 NIH_MGC_21 Homo sapiens CDNA clone IMAGE:3959719 5',
DEFINITION      BP902577
ACCESSION      BP902577
VERSION      BE902577.1 GI:10392908
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 865)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: LCM841 row: D column: 08
High quality sequence stop: 675.
Location/Qualifiers
1. .865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3959719"
/clone_1ib="NIH-MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      204 a      225 c      287 g      149 t


```

COMMENT:
Contact: 1992-93
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 515/2944252
Fax: 515/2942401
Email: cktug91@iastate.edu

FEATURES	Location/Qualifiers
source	1. .557

BASE COUNT	146 a	170 c	137 g	104 t
ORIGIN				

Db 389 GACGCGGTGTAATACTTCCCGCGC 361

SOURCE	human.
ORGANISM	Homo sapiens

JOURNAL
COMMENT

FEATURES
SOURCE

```
/clone_lib="HTB"  
/tissue_type="Hypothalamus"
```

BASE COUNT	185 a	124 c	129 g	190 t	1 others
ORIGIN					

Db 597 GCCGCTGGCGTATAGCAGCCCGGCTGC 629

SOURCE	domestic silkworm.
ORGANISM	<i>Bombyx mori</i>

JOURNAL Unpublished (1999)
COMMENT Contact: Mita K

Email: kmita@nirs.go.jp
PROJECT = 'CREST project by JST'

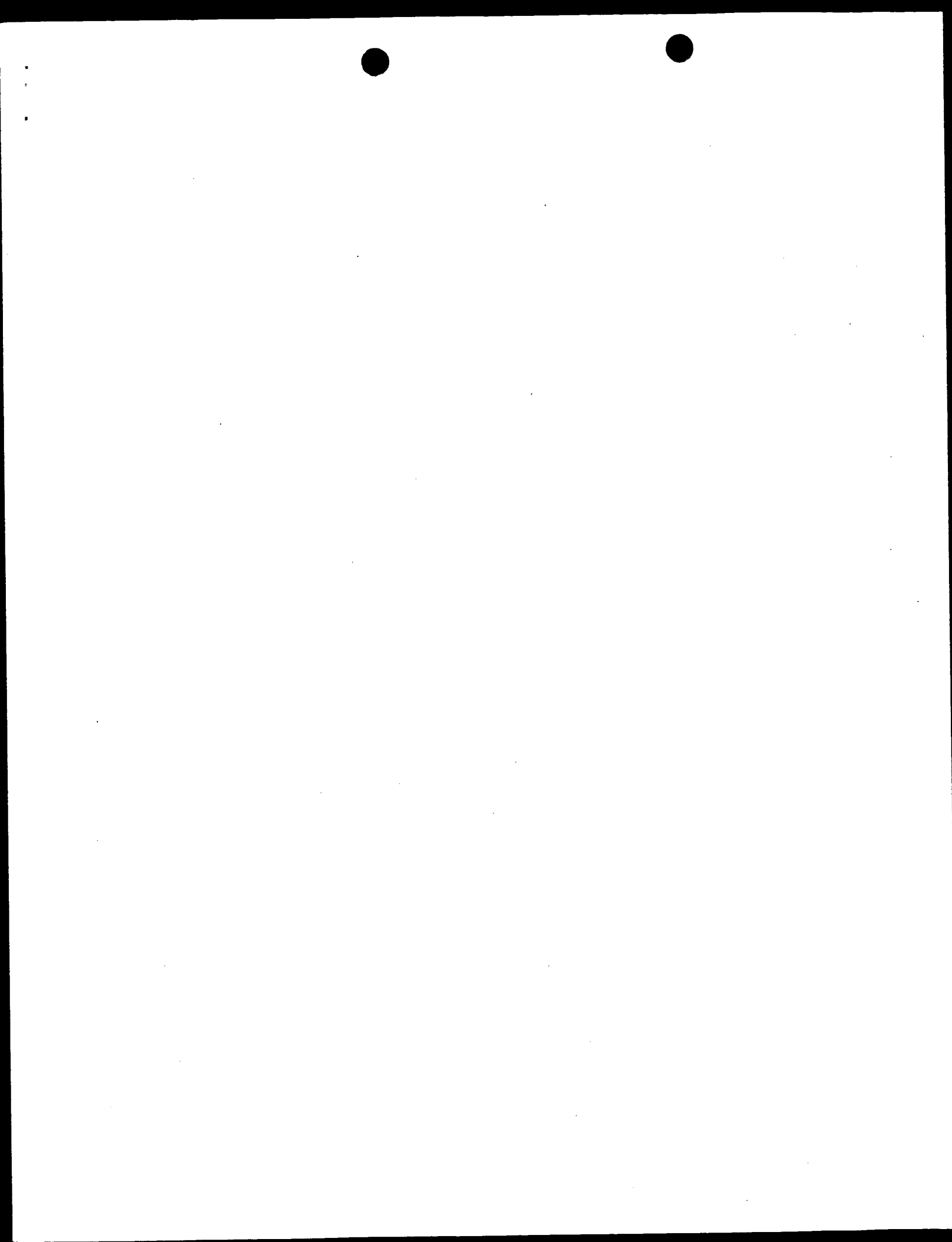
BASE COUNT	173 a	183 c	191 g	166 t	1 others
ORIGIN					

Db 398 AGCCGGGTGTAAGTCGCCGCTCCGG 370

DEFINITION 6024655296F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4593646 5', mRNA sequence.

REFERENCE
1 (bases 1 to 797)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo

Page 5



RESULT 2
 US-09-462-975-1/c
 : Sequence 1, Application US/09462975
 : Patent No. 6303345
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Rohde, Wolfgang
 : APPLICANT: Becker, Dieter
 : APPLICANT: Randles, John W.
 : APPLICANT: Hehn, Alain
 : APPLICANT: Salamini, Francesco
 :
 : TITLE OF INVENTION: USE OF VIRUS DNA AS A PROMOTER

```

1 FILE REFERENCE: 23232.000301
2
3 CURRENT APPLICATION NUMBER: US/09/462,975
4
5 CURRENT FILING DATE: 2000-05-17
6
7 PRIOR APPLICATION NUMBER: PCT/EP99/04345
8
9 PRIOR FILING DATE: 1998-07-13
10
11 PRIOR APPLICATION NUMBER: 19730502.4
12
13 PRIOR FILING DATE: 1997-07-16
14
15 NUMBER OF SEQ ID NOS: 7
16
17 SOFTWARE: FastSeq for Windows Version 4.0
18
19 SEQ ID NO 1
20
21 LENGTH: 1291
22
23 TYPE: DNA
24
25 ORGANISM: Artificial Sequence
26
27 FEATURE:
28
29 OTHER INFORMATION: Description of Artificial Sequence: /note=synthetic
30
31 OTHER INFORMATION: construct
32
33 US-09-462-975-1

```

```

Query Match      100.0%;   Score 31;   DB 4;   length 1291;
Best Local Similarity 100.0%;   Pred. No. 0.00015;
Matches      31;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```

```
OY      1 agccgcggggtataactagccccgcggct   31
        |||||||
Db     70 AGCGGCGGGTAACTAGCCCCGCGGT   40
```

RESULT 3
US-09-462-975-2/c
; Sequence 2, Application US/09462975

```

: GENERAL INFORMATION:
: APPLICANT: Rohde, Wolfgang
: APPLICANT: Becker, Dieter
: APPLICANT: Randles, John W.
: APPLICANT: Hehn, Alain
: APPLICANT: Salami, Francesco
: TITLE OF INVENTION: USE OF VIRUS DNA AS A PROMOTER
: FILE REFERENCE: 2323.0003U1
: CURRENT APPLICATION NUMBER: US/09/462,975
: CURRENT FILING DATE: 2000-05-17
: PRIOR APPLICATION NUMBER: PCT/EP98/04345
: PRIOR FILING DATE: 1996-07-13
: PRIOR APPLICATION NUMBER: 19730502.4
: PRIOR FILING DATE: 1997-07-16
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 31
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: /note=synthetic
: OTHER INFORMATION: construct
US-09-462-975-2

```

Query Match	63.9%	Score 19.8	DB 4	Length 31
Similarity	77.4%	Pred. No. 5.2		
Best Local				
Matches	24	Mismatches	7	Indels 0
				Gaps 0

```

QY      1 agccgcggyggtataactagccccgcgcgt 31
          ||||| | |||||
Db      31 AGCCGGGGGCTAGTATTACCCGGCGCT 1

```

RESULT 4
US-09-462-975-1
Sequence 1, Application US/09462975
; Patent No. 6303345
; GENERAL INFORMATION:
; APPLICANT: Rohde, Wolfgang

```

APPLICANT: Becker, Dieter
APPLICANT: Rangles, John W.
APPLICANT: Hehn, Alain
APPLICANT: Salamini, Francesco
TITLE OF INVENTION: USE OF VIRUS DNA AS A PROMOTER
FILE REFERENCE: 23232.000301
CURRENT APPLICATION NUMBER: US/09/462,975
CURRENT FILING DATE: 2000-05-17
PRIOR APPLICATION NUMBER: PCT/EP98/04345
PRIOR FILING DATE: 1998-07-13
PRIOR APPLICATION NUMBER: 19730502.4
PRIOR FILING DATE: 1997-07-16
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1291
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: /note=synthetic
OTHER INFORMATION: construct
US-09-462-975-1

```

Query Match	63.9%	Score 19.8;	DB 4;	Length 1291;
Best Local Similarity	77.4%;	Pred. No. 6.7;		
Matches	24;	Conservative	0;	Mismatches 7;
			Indels	0;
			Gaps	0

```

QY      1  agccgcgggggtaatactagccccgcgct 31
          |||||  |  |||||
Db     40  agccgcgggggtaatactagccccgcgct 70

```

```

1 5 RESULT
2 US-08-752-760A-1
3 ; Sequence 1, Application US/08752760A
4 ; Patent No. 5877011
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Armentano, Donna
7 ; APPLICANT: Gregory, Richard J.
8 ; APPLICANT: Smith, Alan E.
9 ; TITLE OF INVENTION: CHIMERIC ADENO VIRAL VECTORS
10 ; NUMBER OF SEQUENCES: 3
11 ; CORRESPONDENCE ADDRESS:
12 ; ADDRESSEE: Baker & Botts, L.L.P.
13 ; STREET: 30 Rockefeller Plaza
14 ; CITY: New York
15 ; STATE: NY
16 ; COUNTRY: U.S.A.
17 ; ZIP: 10112
18 ; COMPUTER READABLE FORM:
19 ; MEDIUM TYPE: Diskette
20 ; COMPUTER: IBM Compatible
21 ; OPERATING SYSTEM: DOS
22 ; SOFTWARE: FastSeq Version 2.0
23 ; CURRENT APPLICATION DATA:
24 ; APPLICATION NUMBER: US/08/752,760A
25 ; FILING DATE: 20-NOV-1996
26 ; CLASSIFICATION: 435
27 ; PRIOR APPLICATION DATA:
28 ; APPLICATION NUMBER:
29 ; FILING DATE:
30 ; ATTORNEY/AGENT INFORMATION:
31 ; NAME: Seide, Rochelle K
32 ; REGISTRATION NUMBER: 32,300
33 ; REFERENCE/POCKET NUMBER: A31385
34 ; TELECOMMUNICATION INFORMATION:
35 ; TELEPHONE: 212-705-5000
36 ; TELEFAX: 212-705-5020
37 ; TELEX:
38 ; INFORMATION FOR SEQ ID NO: 1:
39 ; SEQUENCE CHARACTERISTICS:
40 ; LENGTH: 35081 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-752-760A-1

Query Match 60.6%; Score 18.8; DB 2; Length 35081;
Best Local Similarity 76.7%; Pred. No. 22;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 agccgcgggggtaatactagcccccgcgc 30
DB 15096 ACCGCTGGGGTATTACTAGCCACGACG 15125

RESULT 6
US-08-973-068-15/c
Sequence 15, Application US/08973068
Patent No. 6127604
GENERAL INFORMATION:
APPLICANT: Dale, James Langham
APPLICANT: Harding, Robert Maxwell
APPLICANT: Dugdale, Benjamin
APPLICANT: Beetham, Peter Ronald
APPLICANT: Hafner, Gregory John
TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY TOP VIRUS
FILE REFERENCE: 09657/002001
CURRENT APPLICATION NUMBER: US/08/973,068
CURRENT FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: PCT/AU96/00335
EARLIER FILING DATE: 1996-05-31
NUMBER OF SEQ ID NOS: 61
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 60
TYPE: DNA
ORGANISM: Banana Bunchy Top Virus (BBTV)
US-08-973-068-15

Query Match 58.7%; Score 18.2; DB 3; Length 60;
Best Local Similarity 74.2%; Pred. No. 25;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agccgcgggggtaatactagcccccgcgc 31
DB 47 AGCAGGGGGGTAAATAGTCCCGGTGCT 17

RESULT 7
US-08-973-068-12/c
Sequence 12, Application US/08973068
Patent No. 6127604
GENERAL INFORMATION:
APPLICANT: Dale, James Langham
APPLICANT: Harding, Robert Maxwell
APPLICANT: Dugdale, Benjamin
APPLICANT: Beetham, Peter Ronald
APPLICANT: Hafner, Gregory John
APPLICANT: Becker, Douglas Kenneth
TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY TOP VIRUS
FILE REFERENCE: 09657/002001
CURRENT APPLICATION NUMBER: US/08/973,068
CURRENT FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: PCT/AU96/00335
EARLIER FILING DATE: 1996-05-31
NUMBER OF SEQ ID NOS: 61
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 69
TYPE: DNA
ORGANISM: Banana Bunchy Top Virus (BBTV)

US-08-973-068-12

Query Match 58.7%; Score 18.2; DB 3; Length 69;
Best Local Similarity 74.2%; Pred. No. 25;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agccgcgggggtaatactagcccccgcgc 31
DB 56 ACCGCTGGGGTAAATAGTCCCGGTGCT 26

RESULT 8
US-08-973-068-30/c
Sequence 30, Application US/08973068
Patent No. 6127604
GENERAL INFORMATION:
APPLICANT: Dale, James Langham
APPLICANT: Harding, Robert Maxwell
APPLICANT: Dugdale, Benjamin
APPLICANT: Beetham, Peter Ronald
APPLICANT: Hafner, Gregory John
APPLICANT: Becker, Douglas Kenneth
TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY TOP VIRUS
FILE REFERENCE: 09657/002001
CURRENT APPLICATION NUMBER: US/08/973,068
CURRENT FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: PCT/AU96/00335
EARLIER FILING DATE: 1996-05-31
NUMBER OF SEQ ID NOS: 61
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30
LENGTH: 352
TYPE: DNA
ORGANISM: Banana Bunchy Top Virus (BBTV)
US-08-973-068-30

Query Match 58.7%; Score 18.2; DB 3; Length 352;
Best Local Similarity 74.2%; Pred. No. 28;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agccgcgggggtaatactagcccccgcgc 31
DB 104 AGCAGGGGGGTAAATAGTCCCGGTGCT 74

RESULT 9
US-08-973-068-24/c
Sequence 24, Application US/08973068
Patent No. 6127604
GENERAL INFORMATION:
APPLICANT: Dale, James Langham
APPLICANT: Harding, Robert Maxwell
APPLICANT: Dugdale, Benjamin
APPLICANT: Beetham, Peter Ronald
APPLICANT: Hafner, Gregory John
APPLICANT: Becker, Douglas Kenneth
TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY TOP VIRUS
FILE REFERENCE: 09657/002001
CURRENT APPLICATION NUMBER: US/08/973,068
CURRENT FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: PCT/AU96/00335
EARLIER FILING DATE: 1996-05-31
NUMBER OF SEQ ID NOS: 61
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 547
TYPE: DNA
ORGANISM: Banana Bunchy Top Virus (BBTV)
US-08-973-068-24

Query Match 58.7%; Score 18.2; DB 3; Length 547;
Best Local Similarity 74.2%; Pred. No. 29;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agccgcggggtactactagccccgcgcgt 31
DB 366 AGCGCTGGGGGTATATAGTCCCGCAGCGCT 336

RESULT 10
US-08-973-068-29/c

; Sequence 29, Application US/08973068
; Patent No. 6127604

; GENERAL INFORMATION:

; APPLICANT: Dale, James Langham

; APPLICANT: Harding, Robert Maxwell

; APPLICANT: Dugdale, Benjamin

; APPLICANT: Beetham, Peter Ronald

; APPLICANT: Hafner, Gregory John

; APPLICANT: Becker, Douglas Kenneth

; TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY TOP VIRUS

; FILE REFERENCE: 09657/002001

; CURRENT APPLICATION NUMBER: US/08/973,068

; EARLIER FILING DATE: 1998-03-12

; EARLIER APPLICATION NUMBER: PCT/AU96/00335

; EARLIER FILING DATE: 1996-05-31

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 29

; LENGTH: 622

; TYPE: DNA

; ORGANISM: Banana Bunchy Top Virus (BBTV)

US-08-973-068-29

Query Match 58.7%; Score 18.2; DB 3; Length 622;
Best Local Similarity 74.2%; Pred. No. 29;

Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agccgcggggtactactagccccgcgcgt 31
DB 374 AGCAGGGGGGTATATAGTCCCGCTGCT 344

RESULT 11

US-08-973-068-27/c

; Sequence 27, Application US/08973068

; Patent No. 6127604

; GENERAL INFORMATION:

; APPLICANT: Dale, James Langham

; APPLICANT: Harding, Robert Maxwell

; APPLICANT: Dugdale, Benjamin

; APPLICANT: Beetham, Peter Ronald

; APPLICANT: Hafner, Gregory John

; APPLICANT: Becker, Douglas Kenneth

; TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY TOP VIRUS

; FILE REFERENCE: 09657/002001

; CURRENT APPLICATION NUMBER: US/08/973,068

; EARLIER FILING DATE: 1998-03-12

; EARLIER APPLICATION NUMBER: PCT/AU96/00335

; EARLIER FILING DATE: 1996-05-31

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 27

; LENGTH: 624

; TYPE: DNA

; ORGANISM: Banana Bunchy Top Virus (BBTV)

US-08-973-068-27

Query Match 58.7%; Score 18.2; DB 3; Length 624;
Best Local Similarity 74.2%; Pred. No. 29;

Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agccgcggggtactactagccccgcgcgt 31
DB 375 AGCAGGGGGGTATATAGTCCCGCTGCT 345

RESULT 12
US-08-202-186-20/c

; Sequence 20, Application US/08202186

; Patent No. 5756708

; GENERAL INFORMATION:

; APPLICANT: KARAN, Mirko

; APPLICANT: BURNS, Thomas M.

; APPLICANT: DALE, James L.

; APPLICANT: HARDING, Robert M.

; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W.

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/202,186

; FILING DATE: 24-FEB-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: JEFFERY, Donald D.

; REGISTRATION NUMBER: 19,980

; REFERENCE/DOCKET NUMBER: 71611/102 FIKE

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202 672 5300

; TELEFAX: 202 672 5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1075 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: circular

US-08-202-186-20

Query Match 58.7%; Score 18.2; DB 1; Length 1075;
Best Local Similarity 74.2%; Pred. No. 30;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 agccgcggggtactactagccccgcgcgt 31
DB 31 AGCGCTGGGGGTATATAGTCCCGCAGCGCT 1

RESULT 13

US-08-973-068-2/c

; Sequence 2, Application US/08973068

; Patent No. 6127604

; GENERAL INFORMATION:

; APPLICANT: Dale, James Langham

; APPLICANT: Harding, Robert Maxwell

; APPLICANT: Dugdale, Benjamin

; APPLICANT: Beetham, Peter Ronald

; APPLICANT: Hafner, Gregory John

; APPLICANT: Becker, Douglas Kenneth

; TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY TOP VIRUS

; FILE REFERENCE: 09657/002001

; CURRENT APPLICATION NUMBER: US/08/973,068

; CURRENT FILING DATE: 1998-03-12

Wed Jun 26 08:07:42 2002

us-09-462-955b-2.rni

IntelliGenetics

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file SeqeditedvsSeq2inv.res made by tport on Tue Jun 10 14:53:08-PDT.

Query sequence being compared: US-09-462-955B-1-COPY' (1-1288)

Number of scores above cutoff

7

Results of the initial comparison of US-09-462-955B-1-COPY' (1-1288) with:
File : US09462955B.seq

Student	Score
N	100
U	95
M	90
B	85
E	80
S	75
F	70
S	65
E	60
O	55
U	50
E	45
N	40
C	35
E	30

PARAMETERS	
Similarity matrix	Unary
Mismatch penalty	1
Gap penalty	5.00
Gap size penalty	0.33
Cutoff score	1
Randomization group	0
	K-tuple
	Joining
	penalty
	Window
	size
	4
	30
	500

SEARCH STATISTICS		
Mean	Median	Standard Deviation

```

Times:      CPU      Total Elapsed
           00:00:00.00  00:00:00.00

```

Number of residues:	1393
Number of sequences searched:	7
Number of scores above cutoff:	7

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found

The list of best scores is:

Sequence Name	Description	Init.	Opt.	Length	Score	Score	Sig.	Frame
---------------	-------------	-------	------	--------	-------	-------	------	-------

1.	US-09-462-955B-1	Sequence 1,	Application U	1291	86	407	2.19	0
		****	0 standard deviation from mean	****				
2.	US-09-462-955B-2	Sequence 2,	Application U	31	31	0.28	0	

1. US-09-462-955B-1-COPY' (1-1288)
US-09-462-955B-1 Sequence 1, Application US/09462955E

Initial Score	=	86	Optimized Score	=	407	Significance	=	2.19
Residue Identity	=	37%	Matches	=	471	Mismatches	=	743
Gaps	=	43	Conservative Substitutions	=			=	0

10 20 30 40 50 X 60 70
ACCCGTTCCCTGCAGACGAGTCTCTATTCTGTGTCGCGGAACCAACGGCTCGGCTCAGGTGAATCCCTGCA

CTCGAGCACGGAGTCCCCGACTC
X 10 20

80 90 100 110 120 130 140
TTCCAGACAGCTTTCATCCTTGACGTCGCCGACCGGCTTCAAGTGGATGAATCCTTGGAGGTGTCCT

30 40 50 60 70 80 90
GTCCTGGAGTCAAAAAGCCACGATTGGCCCAACGATTTCCTGAGGAACCTGATGTAACTCGCCCTGGAAGACC

150 160 170 180 190 200 210 220
TGGCCAGTTGAAGGTGCAACCTGCTCCCACTATAGCATAGACCAAGTTAAGGCACTCGATGGCCCTTACA

CAAGCGAATACCGAAATGCGTGTGTAACCGAGCTTCGAGG-----AAAGCAAGATGGGCGCGCTG
100 110 120 130 140 150

ACATTCGCCGCTCTTCCTCTGTCTCTAGTTCAGAGTCGAAGCACCACGCGGCGAATGAGAGAACCTTATAGC

160 170 180 190 200 210 220

CAACGGGGGTTAGCAGAGGTTGGGGAGCCCGGGGCTATTACTAGCCCTATTPTTATACTTACGACG

30 240 250 260 270 280 290 300

GGACTTAGCAGAGGTTTTCGCCCCAGATCAACTCAGCCGTCGATCGCGCAAAAAACAATGGACGGTAGAT

Year	310	320	330	340	350	360	370
1980	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1981	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1982	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1983	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1984	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1985	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1986	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1987	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1988	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1989	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1990	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1991	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1992	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1993	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1994	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1995	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1996	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1997	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1998	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1999	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2000	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2001	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2002	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2003	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2004	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2005	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2006	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2007	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2008	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2009	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2010	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2011	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2012	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2013	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2014	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2015	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2016	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2017	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2018	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2019	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2020	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2021	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2022	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2023	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2024	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2025	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2026	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2027	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2028	1.0	1.0	1.0	1.0	1.0	1.0	1.0
2029	1.0	1.0	1.0	1.0	1.0	1.0	1.0

AGAGGCTTGGGGCTTCTCTCCAGCAAGCTACTCACCTGGATTGGTGACGTGGAAAGTGTTCGCTGCACCT

380	390	400	410	420	430	440
530	530	540	550	560	570	580

GCAAGCAAGTACACATTCGTTCTTAATCCAAATGGAMTCTTTATTGTGGGCCCGCCCATTACTGCATTCGA
 | | | | | | | | | | | | | | | | | |
 AAGACACACCCCATTATCCTTGCCGACAATAATCCGAA-CGCCCTTACTATATCTGTCGCCATGATGATCTACT

450	460	470	480	490	500	510
590	600	610	620	630	640	

CGATATTGGCCAAATGCTCTGCTGATATTATTGAAAAATACAGCAGGGGACAGAAATAAAACTGTGGCAATATTTTAAAG
TATIGTCCCGGGCCCTGCGCCGGG-----AGCACCTCCAGGGCTTCGGCTGGGGCAGGAATATCA

520	530	540	550	560	570	580	590	700	710
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

AAAGATGTAACGATACGCGCTGGGCGGCGGATACGCTGTAATTAAGATGACACCAATACCTTAAATATCCCT

|||||
 CCGGGCT
 30

```

TATGTCATCTAATTACACCAATACCGCCGCCGCTATGCTTACATCTTAAATATTCCTGCC
590      600      610      620      630      640      650
720      730      740      750      760      770      780
CAGTTTATTTCTGTCCTGCTGATTTTCAATATATCAGCGACGACATTGCAATATACATGACATG
|||      |||      |||      |||      |||      |||      |||
CAGGCCGAGGCGGGAGGTGCT-----ACCGGCCGAGGCCGGGAAACAATATGATGAGTTATG
660      670      680      690      700      710      720
790      800      810      820      830      840      850
GTCCAGACCCCAAGAT-AACTAAGGGGTCTGATTTGTCCGAACCTAATGCCCTGTCTTACATCTTACA
|||      |||      |||      |||      |||      |||      |||
GGCGGCCCAATAAAGATTCATTGATAGAGCAATCTGTACTTGTCTGACAGTGCACGCAACCA
730      740      750      760      770      780      790
860      870      880      890      900      910      920
GGGCATATTTAATTAACCTTAATTAACCGGTGACA-TCGAGGATTAATTTGTTGGGTCCTGATG
CTTCCACGTCCCAATCCAGGTAGTACTGCTGGAGAAAGAGCCGCAAGCTTATCTACCGTCAATT
800      810      820      830      840      850      860
930      940      950      960      970      980      990
TACGTGTACATTAATTCCTTCTGTTCCACACATGTGTAGAAC--AGTCGGGCTTGATCCGAGATAT
|||      |||      |||      |||      |||      |||      |||
TGTTTTCGGGATCGGACGGCTAGTTGATCTGGCCCAAAAACCTCTGTAAGTCCCTGCTAAGTATAA
870      880      890      900      910      920      930
1000     1010     1020     1030     1040     1050     1060     1070
TTGGCAAAACAGCGACTTCCGTCCTCTCCGTCCTGTCGATATCCAGAGGATTTGTCGCTCCGCTGC
|||      |||      |||      |||      |||      |||      |||
ATACCGCGGGGGCTAGTATTAACCCCGCGCTCCCAACCTCTGTAACCCGCTTGCTATTAATGGGTT
940     950     960     970     980     990     1000     1010
1080     1090     1100     1110     1120     1130     1140     1150
TCTCCGATCCGACAGACACTCAAGCTCAAGTGTGATA-TGGAATGGGAACGGATTTTCAGCGCCCA
|||      |||      |||      |||      |||      |||      |||
CCTCATTCCGCGGCTGCTTCACTGTAACACTAGAGAGAGAGAGAGCGCGGCAATGTTGTAAGCGCA
1020     1030     1040     1050     1060     1070     1080
1150     1160     1170     1180     1190     1200
TCTGTCCAT-----TCCACCGAGCTCCGTCGATCCGATTCCTCGTATCCCGTGGGCTCTCCAGG
|||      |||      |||      |||      |||      |||      |||
TCGAGTCCCTTAACCTGTATGTATAGTTGGGAGAGGTTGCACCTTCAACTGSCCAAGACACCTCC
1090     1100     1110     1120     1130     1140     1150
1210     1220     1230     1240     1250     1260     1270
CGGAGTTCATCAGGTTCTCAGCAAAATCGTGGGCAATGCTGCTTTGACTTCCAGAGAGTGGGACT
|||      |||      |||      |||      |||      |||      |||
AAGATTTCATCACTTGAAGACCGGCTGGGAGCTGCAAGATGAGACTGTTCTTGGAAATGACAGGATTC
1160     1170     1180     1190     1200     1210     1220
1280
1280      X
CCGTGCTCGAG
|||
ACCTGGAGCGGACCGGTGTCGAGACAGACAGATAGACTACTGTTGAGAGGACGGGT
1230     1240     1250     1260     1270     1280
2. US-09-462-955B-1-COPY' (1-1288)
US-09-462-955B-2 Sequence 2, Application US/09462955B
Initial Score = 31 Optimized Score = 31 Significance = 0.28
Residue Identity = 100% Matches = 31 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0
270      280      290      300      310      320      330
GCGGCGAATGAGAGACCTTTAGCAAGCGGGTTAGCAGAGTTGGGAGCGCGGGGTATATATACCC
|||||      |||||||      |||||||      |||||||      |||||||      |||||||
AGCCGCGGGGTATATACTAGCC
X
340      X      360      370      380      390
CCCCGGCTATTATTACTTAGCAGGAGCTTAGCAGAGTTTGGCGCCAGATCAACT

```

The list of best scores is:

Sequence Name	Description	Length	Int. Opt.	Sig.	Frame		
1. US-09-462-955B-1 Sequence 1, Application US/09462955B	**** 2 standard deviations above mean **** US-09-462-955B-1 Sequence 1, Application US/09462955B	1291	1026	1282	2.27 0		
2. US-09-462-955B-2 Sequence 2, Application U	**** 0 standard deviation from mean **** US-09-462-955B-2 Sequence 2, Application U	31	24	24	-0.35 0		
Initial Score = 1026	Optimized Score = 1282	Significance = 2.27					
Residue Identity = 98%	Matches = 1288	Mismatches = 0					
Gaps = 3	Conservative Substitutions = 0						
X	10	20	30	40	50	60	70
CTGCAGCAGCGAGATCCCGACCTGTCCTGAGTCAAAAGCGCCAGGATTTGGCCACCATTTGCTGAGAACCT							
CTGCAGCAGCGAGATCCCGACCTGTCCTGAGTCAAAAGCGCCAGGATTTGGCCACCATTTGCTGAGAACCT							
CTGCAGCAGCGAGATCCCGACCTGTCCTGAGTCAAAAGCGCCAGGATTTGGCCACCATTTGCTGAGAACCT							
X	10	20	30	40	50	60	70
GATGACTCTCCGCTGGAGACCCAGCGGAGATCCGAGATGCGCTTTACACGAGAGCTTGGTGGATGACA							
GATGACTCTCCGCTGGAGACCCAGCGGAGATCCGAGATGCGCTTTACACGAGAGCTTGGTGGATGACA							
GATGACTCTCCGCTGGAGACCCAGCGGAGATCCGAGATGCGCTTTACACGAGAGCTTGGTGGATGACA							
X	10	20	30	40	50	60	70
AGATGGCGCGCTGAAATATCCGCTTCCCATTTCCATATACATTTGAGAGCTTAAATGCTGCTGGATCGGA							
AGATGGCGCGCTGAAATATCCGCTTCCCATTTCCATATACATTTGAGAGCTTAAATGCTGCTGGATCGGA							
AGATGGCGCGCTGAAATATCCGCTTCCCATTTCCATATACATTTGAGAGCTTAAATGCTGCTGGATCGGA							
X	10	20	30	40	50	60	70
GAGCCAGCGGAGATCGACATATCCGCTTCCCATTTCCATATACATTTGAGAGCTTAAATGCTGCTGGATCGG							
GAGCCAGCGGAGATCGACATATCCGCTTCCCATTTCCATATACATTTGAGAGCTTAAATGCTGCTGGATCGG							
GAGCCAGCGGAGATCGACATATCCGCTTCCCATTTCCATATACATTTGAGAGCTTAAATGCTGCTGGATCGG							
X	10	20	30	40	50	60	70
GAATATCTCGAGCTCAAGCCCGACTGGTTCTACATATGTTGGGAAACCAAGAGAGCTATTTGACCAATG							
GAATATCTCGAGCTCAAGCCCGACTGGTTCTACATATGTTGGGAAACCAAGAGAGCTATTTGACCAATG							
GAATATCTCGAGCTCAAGCCCGACTGGTTCTACATATGTTGGGAAACCAAGAGAGCTATTTGACCAATG							
X	10	20	30	40	50	60	70
ATCGAGGAGCCCAAAAGCAAAATTTAATCTCGATGATGATACCCAGGTGTAATTTAGATATTTAAATATGACCTG							
ATCGAGGAGCCCAAAAGCAAAATTTAATCTCGATGATGATACCCAGGTGTAATTTAGATATTTAAATATGACCTG							
ATCGAGGAGCCCAAAAGCAAAATTTAATCTCGATGATGATACCCAGGTGTAATTTAGATATTTAAATATGACCTG							
X	10	20	30	40	50	60	70
TTAGATATGTTGAAGACAGGCGATTCAGTTCGGACAAATATAGAACCCCTTACTTATCTGGTGGATGACCAT							
TTAGATATGTTGAAGACAGGCGATTCAGTTCGGACAAATATAGAACCCCTTACTTATCTGGTGGATGACCAT							
TTAGATATGTTGAAGACAGGCGATTCAGTTCGGACAAATATAGAACCCCTTACTTATCTGGTGGATGACCAT							
X	10	20	30	40	50	60	70
GTGATGATGATCTGATTTTGGCAATGTCCTGCTGATTTTGAATTAATACAGAGAGAGATTTAAATGCTGG							
GTGATGATGATCTGATTTTGGCAATGTCCTGCTGATTTTGAATTAATACAGAGAGAGATTTAAATGCTGG							
GTGATGATGATCTGATTTTGGCAATGTCCTGCTGATTTTGAATTAATACAGAGAGAGATTTAAATGCTGG							
X	10	20	30	40	50	60	70
AAATTTTAAAGTATGTCGATATCAATTAATACCAATATCCGCGCCAGCGGCTATGCTTTACATCATATG							
AAATTTTAAAGTATGTCGATATCAATTAATACCAATATCCGCGCCAGCGGCTATGCTTTACATCATATG							
AAATTTTAAAGTATGTCGATATCAATTAATACCAATATCCGCGCCAGCGGCTATGCTTTACATCATATG							
X	10	20	30	40	50	60	70
AAATATCTGCGAGCGGCAAGGCTGAGAGTCTACCGGCGGAGAGCGGAGCAATATGATATGATGACATTA							
AAATATCTGCGAGCGGCAAGGCTGAGAGTCTACCGGCGGAGAGCGGAGCAATATGATATGATGACATTA							
AAATATCTGCGAGCGGCAAGGCTGAGAGTCTACCGGCGGAGAGCGGAGCAATATGATATGATGACATTA							

Matches	31: Conservative	0: Mismatches	0: Indels	0: Gaps	0:
Qy	1	agccgcggggtatactagcccccgcgcgt	31		
Db	1	AGCCGCGGGGTATAGTACTAGCCCCGCGGCT	31		
RESULT	2				
LOCUS	ARI72770/c	1291 bp	DNA	linear	PAT 17-DEC-2001
DEFINITION	Sequence 1 from patent US 6303345.				
ACCESSION	ARI72770				
VERSION	ARI72770.1	GI:17912261			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1291)				
AUTHORS	Rohde,W., Becker,D., Randles,J.W., Hehn,A. and Salamiini,F.				
TITLE	Use of a virus DNA as promoter				
JOURNAL	Patent: US 6303345-A 1 16-Oct-2001;				
FEATURES	Location/Qualifiers				
Source	1..1291				
BASE COUNT	336 a 323 c 332 g 300 t				
ORIGIN					
Query Match	100.0%; Score 31; DB 6; Length 1291;				
Best Local Similarity	100.0%; Pident. NO. 0.15; 0; Indels 0; Gaps 0;				
Matches	31: Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	agccgcggggtatactagcccccgcgcgt	31		
Db	70	AGCCGCGGGGTATAGTACTAGCCCCGCGGCT	40		
RESULT	3				
LOCUS	CFDGC	1291 bp ss-DNA			circular VRL 02-AUG-1993
DEFINITION	Cocount foliar decay virus, complete genome.				
ACCESSION	M29963				
VERSION	M29963.1	GI:323306			
KEYWORDS	circular: complete genome.				
ORIGIN	Cocount foliar decay virus DNA, clones NMORG, PORC, and PORGSNU3A.				
ORGANISM	Cocount foliar decay virus				
REFERENCE	1 (bases 1 to 1291)				
AUTHORS	Rohde,W., Randles,J.W., Langridge,P. and Hanold,D.				
TITLE	Nucleotide sequence of a circular single-stranded DNA associated with cocount foliar decay virus				
JOURNAL	Virology 176, 648-651 (1990)				
COMMENT	90266484				
MEDLINE					
FEATURES	Draft entry and printed sequence for [1] kindly submitted by W.Rhode, 15-MAR-1989, for release after publication.				
Source	Location/Qualifiers				
1..1291					
/organism="Cocount foliar decay virus"					
/db_xref="taxon:12474"					
40..70					
/note="stem-loop structure"					
103..975					
/note="ORF 1"					
/codon_start=1					
/protein_id="AAA42894.1"					
/db_xref="GI:323307"					
/translation="WGSSTRMCTFLNLTETEEBANVYRTESLNTLYATIGDVAES					
TCGRHLQGIHILKTGRGLGTIVGNDRIHLEPTRGSDEONRDCSEKRYLLHGCVR					
TPQVGRPRILAGFAEEDDELIEDPGGYRRGVGASVETIRMAAENPFPFYHNMQ					
LEVLAIGRPADDRFTILWICGRDGDGSVAKYLGILKPDWFTYCGGTRKDVLYOLE					
DKRNIILDPKRCNLEYINVALLECVKNRARSDDKEPLSYLGEFHVHVAIVANLP					
YKIKSRDKIKLMNT"					
314..775					

FEATURES	LOCATION/Qualifiers
CDs	<p> <code>/note="ORF 2"</code> <code>/codon_start=1</code> <code>/protein_id="AA042895.1"</code> <code>/db_xref="GI:323308"</code> <code>/translation="MTGFTWSRPVPTNRIETTVARNCCFSSSTESRLVLESKGMDPN DLNNLNNSAMKTAQDTEDALYTELRWNQSDPLKINSHFITIGSKCLRESQRT IAOSSGVADEETGSPCLPNISDSPTGSHVPEPFTYTSRSTONE"</code> <code>complement(422..568)</code> <code>/note="ORF 6"</code> <code>/note="ORF 6"</code> <code>/codon_start=1</code> <code>/protein_id="AA042896.1"</code> <code>/db_xref="GI:323309"</code> <code>/translation="MEMGTDFOFDIISIPPKLRVDRIRGIRGLPGVGHVYPOQIYGP IVAF"</code> <code>639..797</code> <code>/note="ORF 3"</code> <code>/codon_start=1</code> <code>/protein_id="AA042897.1"</code> <code>/db_xref="GI:323310"</code> <code>/translation="MRTRRRRRREVRVCOISRTQARLVLMHMMNOKGIVPVHNGPKTK FNPRCTQY"</code> <code>complement(823..987)</code> </p>
CDs	<p> <code>/note="ORF 5"</code> <code>/codon_start=1</code> <code>/protein_id="AA042898.1"</code> <code>/db_xref="GI:323311"</code> <code>/translation="MRTNLNPQFYSVPADFQIIRQDICKYEYHMHVPEPKITKGFVFL RTECPVLTNF"</code> <code>1098..1286</code> <code>/note="ORF 4"</code> <code>/codon_start=1</code> <code>/protein_id="AA042899.1"</code> <code>/db_xref="GI:323312"</code> <code>/translation="MNRVWGSGPTIKDSIWIRNLCLICQCTOPLSTSPLOYSLLEKK AASLYPSICFCACGRLS"</code> <code>336 a 323 c 332 g 300 t</code> </p>
BASE COUNT	336 a 323 c 332 g 300 t
ORIGIN	
Query Match	100.0%; Score 31; DB 14; Length 1291;
Best Local Similarity	100.0%; Pred. No. 0.15;
Matches 31; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 agccgcg9gggtactactagcccccgcgcgt 31
Db	70 AGCCGCGGGGTAATACTAGCCCCCGCGCT 40
RESULT 4	
AE006200/c	9294 bp DNA linear BCT 08-MAR-2001
LOCUS	
DEFINITION	Pasteurella multocida pm70 section 167 of 204 of the complete genome.
ACCESSION	AE006200 AE004439
VERSION	AE006200.1 GI:12722026
KEYWORDS	
SOURCE	Pasteurella multocida.
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Pasteurella.
REFERENCE	1 (bases 1 to 9294)
AUTHORS	May,B.,J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and Kapur,V.
TITLE	Complete genomic sequence of Pasteurella multocida, pm70
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
MEDLINE	21145866
REFERENCE	2 (bases 1 to 9294)
AUTHORS	Zhang,Q. and Kapur,V.
TITLE	Direct Submission
JOURNAL	Submitted (24-OCT-2000) Department of Veterinary Pathobiology, University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN 55108, USA

source
1. .9294
/organism="Pasteurella multocida"
/strain="PM70"
/db_xref="taxon:747"
complement(161. .853)
gene
/gene="PM1626"
complement(161. .853)
CDS
/gene="PM1626"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03710.1"
/db_xref="GI:12722027"
/translation="MKTGLIGGMSPESTVYVYRLINLQAKSGNSAPLVYVNE
FBOIVCLOGEMKAGETIATAKKLQIGAGGILLATNTMKVAPQIIDVDPVFL
HIDATAKAIKINLTQVALLGTOFTMODPFYRQDOLLHGGITPLVNPQKQOETHRII
FTELCVGQIKPEKQOFYLDVINDLAKGAGVILGCTEGILLINQDNLTFEPDTTLL
HQAQAEEFLITE"
gene
complement(872. .1102)
/gene="11VM"
/note="PM1627"
complement(872. .1102)
CDS
/gene="11VM"
/codon_start=1
/transl_table=11
/product="11VM"
/protein_id="AAK03711.1"
/db_xref="GI:12722028"
/translation="MOTYQFLDVNQRPETLERILRIHRHGFYVVAAMNTAQEQAN
IDFYKSDRAEFALLTNLSKYDYDLAIHQVRA"
gene
complement(1140. .2789)
/gene="11VG"
/note="PM1628"
complement(1140. .2789)
CDS
/gene="11VG"
/codon_start=1
/transl_table=11
/product="11VG"
/protein_id="AAK03712.1"
/db_xref="GI:12722029"
/translation="MNGARLVTESLKTGHVTVTVYVYVGAIMPYDALYDAGIEHLIC
RNEQGAAMAAGVARSQGVCIATSGPGAAALITGIDALMDSIPVAVITGVVASP
LIGTDAFQEDVGLSLACTKHSFLVSGIEELPVYAKAFETIAGSRGPGPLIDIPD
IOIADTVTPETYSVAKPTLQSOAHLIDALTLIKOAKRPVLYVGGGVMGAVAYALAE
FLAYOMPVSSTLKGISVHPNDPYMGIMGHTKXANVYVOCEDLLTAFGARFDR
VTGKIDPEAPAKYTHADIDMAEIGKLRPVYALCGDIVAFNLSTPLNTAEQAH
KOLQKTHDFCTSTNOGDTFIMPLWLNLSKRPQSAVITTDVGQHQMSAHOQHTA
PENYITSGYTFMGFLPAALGAKARAPDEVILISGDSIMANNIOELGTFKRGKTPV
KIVLLDNORLQMVQOELFENARFSMTILDNDPFYMLAAAFGIOGERITKGEVEG
ALDRILASKEAYFLHVCISSNENWPLVPPNACMLDWVEQW"
gene
3149. .3967
/gene="PM1629"
3149. .3967
CDS
/gene="PM1629"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03713.1"
/db_xref="GI:12722030"
/translation="MSFFPRRAVYSDLTLLNANHMIIGPTIOTLQQLAAKIDIML
ATGRNHTDILPILKKVINIEKAVMTTSGARAQDQGMILVADLPEDIAFINLLPDD
RQRCVSAVQGDWFINIDVQLRKYHODSGFMVEYVDFAOHNGRETEKFFIGREPO
DLIGLEOHLQANADTISITYSEVLCLEIMNNKVSASALEKVLADRDYDLOHCLAFG
DGMNDVOMLSRVKGCYMGNAADPLKRCACPHLEVLIGNNAQESVANYIRTYFDIE"
gene
3983. .4387
/gene="CrcB"
/note="PM1630"
3983. .4387
CDS
/gene="CrcB"
/codon_start=1
/transl_table=11
/product="CrcB"
/protein_id="AAK03714.1"
/db_xref="GI:12722031"
/translation="MAGSRVYSVGGDIIITISGALGALSRLGILLNPLFSFSL
GITLVANLIGLITIGVFLAFWQYQCSAEMRLFEVGLSITTFSTFSAEVIENTLIO
QKVLAGLMLASGHLIGLITFALGVFWRWQ"
gene
complement(4365. .5318)
/gene="P1dAb"
/note="PM1631"
complement(4365. .5318)
CDS
/gene="P1dAb"
/codon_start=1
/transl_table=11
/product="P1dAb"
/protein_id="AAK03715.1"
/db_xref="GI:12722032"
/translation="MIEPFTHEPALAEHPRAOPPLDQVGGKKGAVITRYEVODN
FVQNLVILVNGRENLKWTETELAYDYEQGGYVULVDRHGQGYSORLLPDHKGYID
EFRFVADDMAILLAKISALPYENQHLIAGLGLISTYLYANDQVSAVSAPFY
G1PLKHSFRDELINLMMLDQGSRYVFGKRYOPADLDNDSCCRRTRMWRNIR
RYPIHLGCTPFRVHLGFSAIKKLPTILPRIEIPVILLOSEKEKIVEKNNLKRLTAL
LPRELTOIQAOSKHEILFERDLIRSEALKKISQFTANSK"
gene
complement(5460. .6569)
/gene="Asd"
/note="PM1632"
complement(5460. .6569)
CDS
/gene="Asd"
/codon_start=1
/transl_table=11
/product="Asd"
/protein_id="AAK03716.1"
/db_xref="GI:12722033"
/translation="MONVFTIGRWGVSYLMDROMOENDFANINPFFVTSQAGNA
PVFGKQAGELKDAFDIAELOKDIITVTCOGGYTNEVPKLRAGMKGYWIDAASAL
RMEKDAIIVLDPVNONVISEGLKNGIKTYVGGNCTVSLMMLGGLFEEDIVENYVA
TYOASGAGKKNMRELLVOMGLEDGSKYDLDNPASTIIDIERKTAARIGLPIENF
GAPLAGSLIPWIKLMEDGQTKREMGYAFNTKLTGLDSPITPDSGLCVRIGALRHS
QATFIRKKDLPLEITEOILLASHNEWKVIPNEKDVLTRELPPTKVTGLSVVGRLR
KLKMGPEVLAFTVGDQLWGAEPILRILKOLVA"
gene
6770. .7009
/gene="PM1633"
6770. .7009
CDS
/gene="PM1633"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03717.1"
/db_xref="GI:12722034"
/translation="MDLGFYFAPYRDERYOKRSYVITQPTERKLKASISLRNPSD
SDNESKARNTIDINIKOITBENKIPKRSIELFQ"
gene
7028. .7381
/gene="PM1634"
7028. .7381
CDS
/gene="PM1634"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03718.1"
/db_xref="GI:12722035"
/translation="MNPNMKKILKESKGLPFLALAEFDYFGSVKSDSETKELFS
FVAELMDGELKFAIRGKFLGSIIEQDILFRQAMPDHYDDELRYDIDHLMWITYAP
AGAVWICEDGYEFW"
gene
7583. .7948
/gene="PM1635"
7583. .7948
CDS
/gene="PM1635"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03719.1"
/db_xref="GI:12722036"
/translation="MLEKSEAYRKRIEENNGVLDELIFANFOCYCQDIDRDKPEIKE
MFLMDADLMKGTGLATKGFLEGSIEEQDILFRQAMPDHYDDELRYDIDHLMWIT
TYAPAGAVWIMEDGYEDWT"
gene
8372. .8637

```

/gene="PM1636"
8272. .8637
/gene="PM1636"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03720.1"
/db_xref="GI:12722037"
/translation="MLEKSEAOYKRRIVENNRGLTIIGMFSSFOVCYGDIRDPEIKE
MPELEIMTDIMKTGTGKLTATGKGFLEGSIEQIDLFQRAWPDHYDKLELEYDINIMWH
VYPACAVNICEBGEYEWMT"
8707. .9144
/gene="PM1637"
8707. .9144
/gene="PM1637"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03721.1"
/db_xref="GI:12722038"
/translation="TMLYRRWSSRSRSPDERYORDKEVNMLEKSEAOYKRRIVENNRGL
ALDLPFASFOVCYGDIRDPEIKEPELEIMADIKTELKLTATGKGFLEGSIEQIDIV
FQRAWPDHYDKLELEYDINIMWHVYVYPACAVNIMWEDGEYEWMT"

BASE COUNT      2772 a      1733 c      1979 g      2810 t

ORIGIN
Query Match      68.4%; Score 21.2; DB 1; Length 9294;
Best Local Similarity 88.5%; Pred. No. 4e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY      5 gcggggggtactactagcccccgcggc 30
||||| ||||||| |||||||
Db      676 GCGGGGGAATACGTAGCCACTCGGC 651
||||| ||||||| |||||||

RESULT 5
AB000926/c
LOCUS DEFINITION      981 bp      DNA      circular VRL 30-OCT-1998
AB000926
Milk vetch dwarf virus genome segment 7 encoding viral
non-structural protein, complete sequence.
ACCESSION      AB000926
VERSION      AB000926.1 GI:3808185
KEYWORDS      viral non-structural protein.
SOURCE      milk vetch dwarf virus (lab_host:Pisum sativum) DNA.
ORGANISM      Viruses; ssDNA viruses; Nanovirus.
1 (bases 1 to 981)
Sano, Y.
Direct Submission
Submitted (08-FEB-1997) Yoshitaka Sano, Kyoto Institute of
Technology, Department of Applied Biology; Matsugasaki, Sakyo-ku,
Kyoto 606-8585, Japan (Tel:075-724-7764, Fax:075-724-7764)
2 (sites)
Sano, Y., Wada, M., Hashimoto, Y., Matsumoto, T. and Kojima, M.
Sequences of ten circular ssDNA components associated with the milkvetch dwarf virus genome
J. Gen. Virol. 79 (Pt 12), 3111-3118 (1998)
99094638
On Oct 29, 1998 this sequence version replaced gi:3798663.
FEATURES
Location/Qualifiers
1. .981
/organism="milk vetch dwarf virus"
/vision
/db_xref="taxon:67585"
/chromosome="segment 7"
/lab_host="Pisum sativum"
6. .34
280. .286
346. .723
/function="unknown"
/note="putative"
/codon_start=1

```

BASE COUNT	291 a	170 c	214 g	306 t
ORIGIN	polyA_signal			
Query Match	Best Local Similarity	85.2%;	Score 20.6;	DB 14; Length 981;
Matches	23; Conservative	0;	Mismatches	4; Indels 0; Gaps 0;
Db	31	GGGGGGTAATCTAAGCCCGCCCT	5	
QY	5	gcggggtaatactagcccccgcgcgc	31	
LOCUS	AB000924/c	non-structural protein, complete sequence.		
DEFINITION	AB000924	989 bp	DNA	circular VRL 30-OCT-1998
ACCESSION	AB000924	GI:3808183		
VERSION	AB000924.1	GI:3808183		
KEYWORDS	viral non-structural protein.			
SOURCE	milk vetch dwarf virus (lab_host:Pisum sativum) DNA.			
ORGANISM	milk vetch dwarf virus			
REFERENCE	1 (bases 1 to 989)			
AUTHORS	Sano, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-FEB-1997) Yoshitaka Sano, Kyoto Institute of Technology, Department of Applied Biology; Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan (Tel:075-724-7764, Fax:075-724-7764)			
AUTHORS	2 (sites)			
TITLE	Sano, Y., Wada, M., Hashimoto, Y., Matsumoto, T. and Kojima, M.			
JOURNAL	Sequences of ten circular ssDNA components associated with the milk			
MEDLINE	J. Gen. Virol. 79 (Pt 12), 3111-3118 (1998)			
COMMENT	On Oct 29, 1998 this sequence version replaced gi:3798659.			
FEATURES	Sequence updated (29-Nov-1997).			
Source	Location/Qualifiers			
	1..989			
	/organism="milk vetch dwarf virus"			
	/virion			
	/db_xref="taxon:67585"			
	/chromosome="5"			
	/lab_host="Pisum sativum"			
	6..34			
	279..285			
	376..628			
	/function="unknown"			
	/note="putative"			
	/codon_start=1			
	/product="viral non-structural protein"			
	/protein_id="BA033984.1"			
	/db_xref="GI:3798660"			
	/translation="MSVNDYSLFPEADDELVSSEKRLIANDCHDDSDQIVNKKEDLFL			
	CDMSKVVAKLQFRICVKKRIIDITLLGCMKKVYTELKTSPSEKSLSILOKKNNIIS			
	DGNVLTIGRMFFININOLINCKMKTIRIEDVYPICITLVHNNPEVIDI"			
BASE COUNT	284 a	167 c	210 g	328 t
ORIGIN	polyA_signal			
Query Match	Best Local Similarity	85.2%;	Score 20.6;	DB 14; Length 989;
Matches	23; Conservative	0;	Mismatches	4; Indels 0; Gaps 0;

QY 5 gccggggtataactagcccccgcgcgt 31
 DB 31 gccggggtataactagcccccgcgcgt 5

RESULT 7
 LOCUS AB000923/c 990 bp DNA circular VRL 30-OCT-1998
 DEFINITION Milk vetch dwarf virus genome segment 4 encoding viral non-structural protein, complete sequence.
 ACCESSION AB000923
 VERSION AB000923.1 GI:3808182
 KEYWORDS viral non-structural protein.
 SOURCE milk vetch dwarf virus (lab_host:Pisum sativum) DNA.
 ORGANISM Viruses; ssDNA viruses; Nanovirus.
 REFERENCE 1 (bases 1 to 990)
 AUTHORS Sano, Y.
 JOURNAL Direct Submission
 TITLE Submitted (08-FEB-1997) Yoshitaka Sano, Kyoto Institute of Technology, Department of Applied Biology; Matsugasaki, Sakyo-Ku, Kyoto 606-8585, Japan (Tel:075-724-7764, Fax:075-724-7764)
 REFERENCE 2 (sites)
 AUTHORS Sano, Y., Wada, M., Hashimoto, Y., Matsumoto, T., and Kojima, M.
 TITLE Sequences of ten circular ssDNA components associated with the milk vetch dwarf virus genome
 JOURNAL J. Gen. Virol. 79 (pt 12), 3111-3118 (1998)
 COMMENT On Oct 29, 1998 this sequence version replaced gi:3798657.
 FEATURES
 source
 1..990
 location/Qualifiers
 /organism="milk vetch dwarf virus"
 /vtilon
 /db_xref="taxon:67585"
 /chromosome="segment 4"
 /lab_host="Pisum sativum"
 6..34
 279..285
 319..828
 /function="unknown"
 /codon_start=1
 /product="viral non-structural protein"
 /protein_id="BA33983.1"
 /db_xref="GI:3798658"
 /translation="MGKRYAHLPLEIREIVDHOERKKKEFLKALDSCRRHEA
 LIIDPSPALNSIKFLIALSDYVGNQFTPLIRKKKDPSPKVFGEDEQHHLY
 GSDMDLSCGELFIPDEDDLTVEGVIVRCSDLQLFKSLGIIVIVSKHCJWA
 PLSEIVIK"

polyA_signal 820..825
 BASE COUNT 285 a 161 c 226 g 318 t
 ORIGIN

Query Match 66.5%; Score 20.6; DB 14; Length 990;
 Best Local Similarity 85.2%; Pred. No. 1.1e+03;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 gccggggtataactagcccccgcgcgt 31
 DB 31 gccggggtataactagcccccgcgcgt 5

RESULT 8
 LOCUS AC098227/c 100542 bp DNA linear HTG 20-DEC-2001
 DEFINITION Rattus norvegicus clone CH230-58J19, *** SEQUENCING IN PROGRESS
 ACCESSION AC098227
 VERSION AC098227.3 GI:17974408
 KEYWORDS HTG; HTGS; PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 100542)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Alstbrooks, S.L., Amaral, H.C., Are, J.R., Banks, T., Barbieri, J.,
 Benton, J., Binage, K., Blankenburg, R., Bonin, D., Bouck, J.,
 Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
 Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
 Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
 Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
 Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
 Denny, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
 Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C.,
 Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
 Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
 Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S.,
 Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
 Hernandez, J., Hernandez, O., Hodgeson, A., Hogues, M., Hollway, C.,
 Hollins, B., Honsi, F., Howard, S., Huber, J., Huliy, S., Hume, J.,
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
 Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W.,
 Louisedge, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
 Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,
 Martinez, E., Massey, E., Mashiney, E., McLeod, M.P., Meador, B.,
 Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
 Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwenwo, S.,
 Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
 Peery, J., Perez, L., Peters, L., Pickens, R., Prims, E., Pu, L.L.,
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojupokan, I., Rolle, M.,
 Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N.,
 Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,
 Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K.,
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
 Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S.,
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 100542)
 Worley, K.C.
 Direct Submission
 Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Dec 20, 2001 this sequence version replaced gi:17062637.

COMMENT

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project information
 Center project name: GTVS
 Center clone name: CH230-58J19
 Summary Statistics
 Assembly program: Phrap; version 0.990329first call to
 findphraplist
 Consensus quality: 85818 bases at least Q40
 Consensus quality: 94218 bases at least Q30
 Consensus quality: 99376 bases at least Q20
 Estimated insert size: 54170; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-1p estimation
 Quality coverage: 0.6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 59 contigs. The true order of the pieces

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 * 4312: contig of 4312 bp in length
 * 4313
 * 4412: gap of unknown length
 * 4413
 * 7638: contig of 3226 bp in length
 * 7639
 * 7739
 * 10359: contig of 2621 bp in length
 * 10360
 * 10459: gap of unknown length
 * 10460
 * 13662: contig of 3203 bp in length
 * 13663
 * 13762: gap of unknown length
 * 13763
 * 15839: contig of 2077 bp in length
 * 15840
 * 17655: gap of unknown length
 * 17656
 * 17755: contig of 1716 bp in length
 * 17756
 * 19386: contig of 1631 bp in length
 * 19387
 * 19486: gap of unknown length
 * 19487
 * 21390: contig of 1904 bp in length
 * 21391
 * 21490: gap of unknown length
 * 23084: contig of 1594 bp in length
 * 23085
 * 23184: gap of unknown length
 * 25329: contig of 2145 bp in length
 * 25330
 * 25429: gap of unknown length
 * 25430
 * 26940: contig of 1511 bp in length
 * 26941
 * 27040: gap of unknown length
 * 27041
 * 28658: contig of 1618 bp in length
 * 28659
 * 28758: gap of unknown length
 * 30808: contig of 2050 bp in length
 * 30809
 * 30908: gap of unknown length
 * 32751: contig of 1843 bp in length
 * 32752
 * 32851: gap of unknown length
 * 32852
 * 33882: contig of 1041 bp in length
 * 33893
 * 33992: gap of unknown length
 * 33993
 * 35243: contig of 1251 bp in length
 * 35244
 * 35343: gap of unknown length
 * 35344
 * 36869: contig of 1526 bp in length
 * 36870
 * 36970: gap of unknown length
 * 38670
 * 38969: contig of 2000 bp in length
 * 38970
 * 39059: gap of unknown length
 * 40571: contig of 1502 bp in length
 * 40572
 * 40671: gap of unknown length
 * 40672
 * 41748: contig of 1077 bp in length
 * 41749
 * 41848: gap of unknown length
 * 41849
 * 43764: contig of 1916 bp in length
 * 43765
 * 43864: gap of unknown length
 * 43865
 * 44951: contig of 1087 bp in length
 * 44952
 * 45051: gap of unknown length
 * 45052
 * 47114: contig of 2063 bp in length
 * 47115
 * 47214: gap of unknown length
 * 47215
 * 48401: contig of 1187 bp in length
 * 48402
 * 48501: gap of unknown length
 * 48502
 * 49778: contig of 1277 bp in length
 * 49779
 * 49878: gap of unknown length
 * 49879
 * 51538: contig of 1660 bp in length
 * 51539
 * 51638: gap of unknown length
 * 51639
 * 53492: contig of 1854 bp in length
 * 53493
 * 53592: gap of unknown length
 * 53593
 * 54753: contig of 1161 bp in length
 * 54754
 * 54853: gap of unknown length
 * 54854
 * 56334: contig of 1481 bp in length
 * 56335
 * 56434: gap of unknown length
 * 56435
 * 57728: contig of 1294 bp in length
 * 57729
 * 57828: gap of unknown length
 * 57829
 * 59995: contig of 2167 bp in length
 * 59996
 * 60095: gap of unknown length
 * 60096
 * 61595: contig of 1500 bp in length
 * 61596
 * 61696
 * 63519: contig of 1824 bp in length
 * 63520
 * 63619: gap of unknown length
 * 63620
 * 64853: contig of 1234 bp in length

* 64854 64953: gap of unknown length
 * 64954 66047: contig of 1094 bp in length
 * 66048 66147: gap of unknown length
 * 66148 67853: contig of 1706 bp in length
 * 67854 67953: gap of unknown length
 * 67954 69623: contig of 1670 bp in length
 * 69624 69723: gap of unknown length
 * 69724 71072: contig of 1349 bp in length
 * 71073 71172: gap of unknown length
 * 71173 72453: contig of 1281 bp in length
 * 72454 72553: gap of unknown length
 * 72554 73633: contig of 1060 bp in length
 * 73634 73733: gap of unknown length
 * 73734 74776: contig of 1043 bp in length
 * 74777 74877: gap of unknown length
 * 74878 76684: contig of 1808 bp in length
 * 76685 76785: gap of unknown length
 * 76786 78396: contig of 1611 bp in length
 * 78397 78496: gap of unknown length
 * 78497 79759: contig of 1264 bp in length
 * 79760 79859: gap of unknown length
 * 79860 80989: contig of 1129 bp in length
 * 80989 81088: gap of unknown length
 * 81089 82146: contig of 1058 bp in length
 * 82147 82247: gap of unknown length
 * 82248 83416: contig of 1170 bp in length
 * 83417 83516: gap of unknown length
 * 83517 84881: contig of 1365 bp in length
 * 84882 84981: gap of unknown length
 * 84982 86401: contig of 1420 bp in length
 * 86402 86501: gap of unknown length
 * 86502 87826: contig of 1325 bp in length
 * 87827 87926: gap of unknown length
 * 87927 89502: contig of 1576 bp in length
 * 89503 89602: gap of unknown length
 * 89603 90937: contig of 1335 bp in length
 * 90938 91037: gap of unknown length
 * 91038 92578: contig of 1541 bp in length
 * 92579 92678: gap of unknown length
 * 92679 94135: gap of unknown length
 * 94136 94234: contig of 1456 bp in length
 * 94235 95305: contig of 1071 bp in length
 * 95306 95405: gap of unknown length
 * 95406 97064: contig of 1659 bp in length
 * 97065 97164: gap of unknown length

Query Match 65.8%; Score 20.4; DB 2; Length 100542;
 Best Local Similarity 80.0%; Pred. No. 4.6e+02;
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 gccgcgggggtatctatcccccgcgcgt 31
 DB 75204 gccgcgggggtatctatcccccgcgcgt 75175

RESULT 9

AC098923 179030 bp DNA linear HTG 20-DEC-2001
 LOCUS Rattus norvegicus clone CH230-154B15, *** SEQUENCING IN PROGRESS
 DEFINITION *** 68 unordered pieces.

AC098923
 AC098923.4 GI:17973852
 VERSION
 KEYWORDS
 HTG: HTGS, PHASEL.
 SOURCE
 ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 179030)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Alstrooms, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbara, J.,
 Benton, J., Blum, K., Blankenship, K., Bonnin, D., Bouck, J.,
 Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Butay, C.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Burch, P., Burrell, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G.,
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Dem, A.L., Ding, Y., Dinh, H.H., Douthett, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, C.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garcia, N., Gill, R., Correll, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,
Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Humm, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J.,
Kovar, C., Kralovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtenberg, O., Lieu, C., Liu, J., Liu, W.,
Loui, S., Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,
Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabat, K.,
Moran, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, J.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nockem, S.,
Ogutu, M., Okunolu, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojuckan, I., Roife, M.,
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N.,
Stinson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, V., Telford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Washington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S.,
Wotley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,
Weinstock, G. and Gdbbs, R.

Direct Submission
2 (bases 1 to 179030)
Worley, K.C.

Submitted (06-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gl:17064677.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G10K
Center clone name: CH230-154B15
----- Summary Statistics
Assembly program: Phrap; version 0.990329first call to
findPhrapList
Consensus quality: 137096 bases at least Q40
Consensus quality: 145080 bases at least Q30
Consensus quality: 152950 bases at least Q20
Estimated insert size: 138980; sum-of-coverage estimation
Quality coverage: 0x in Q20 bases; agarose-tp estimation
Quality coverage: 2.2x in Q20 bases; sum-of-coverage estimation

----- NOTE: Estimated insert size may differ from sequence length
* (see <http://www.hgsc.bcm.tmc.edu/docs/genbank-draft-data.html>).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 68 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved. 7467: contig of 7467 bp in length

1
7467: contig of 7467 bp in length

7468 7567: gap of unknown length
7568 12115: contig of 4548 bp in length
7569 12115: gap of unknown length
12216 12215: contig of 8098 bp in length
12217 20314: gap of unknown length
20314 20413: contig of 5176 bp in length
25589 25590: gap of unknown length
25590 25690: contig of 5745 bp in length
31435 31534: gap of unknown length
31535 37595: contig of 6061 bp in length
37596 42172: contig of 4477 bp in length
42173 42272: gap of unknown length
42273 46334: contig of 3862 bp in length
46334 50605: gap of unknown length
50605 50705: gap of unknown length
50705 55202: contig of 4497 bp in length
55202 55303: gap of unknown length
55303 58407: contig of 3105 bp in length
58407 58508: gap of unknown length
58508 61287: contig of 2779 bp in length
61287 61386: gap of unknown length
61386 64467: contig of 3081 bp in length
64467 64567: gap of unknown length
64567 68498: contig of 3931 bp in length
68498 72449: gap of unknown length
72449 73549: contig of 4851 bp in length
73549 73550: gap of unknown length
73550 77402: contig of 3853 bp in length
77402 77503: gap of unknown length
77503 81027: contig of 3525 bp in length
81027 81128: gap of unknown length
81128 84144: contig of 3017 bp in length
84144 86244: gap of unknown length
86244 86680: contig of 2436 bp in length
86680 86780: gap of unknown length
86780 89210: contig of 2430 bp in length
89210 89310: gap of unknown length
89310 92339: contig of 3029 bp in length
92339 92439: gap of unknown length
92439 94951: contig of 2512 bp in length
94951 95051: gap of unknown length
95051 95052: gap of unknown length
95052 98586: contig of 3535 bp in length
98586 98687: gap of unknown length
98687 100597: contig of 1911 bp in length
100597 100697: gap of unknown length
100697 102766: contig of 2066 bp in length
102766 102866: gap of unknown length
102866 105237: contig of 2371 bp in length
105237 105337: gap of unknown length
105337 107305: contig of 1966 bp in length
107305 107405: gap of unknown length
107405 109057: contig of 1652 bp in length
109057 109157: gap of unknown length
109157 111454: contig of 2297 bp in length
111454 111554: gap of unknown length
111554 113704: contig of 2150 bp in length
113704 113804: gap of unknown length
113804 115933: contig of 2129 bp in length
115933 116033: gap of unknown length
116033 118193: contig of 2160 bp in length
118193 118293: gap of unknown length
118293 120414: contig of 2121 bp in length
120414 120514: gap of unknown length
120514 122311: contig of 1797 bp in length
122311 122411: gap of unknown length
122411 125628: contig of 3217 bp in length
125628 125728: gap of unknown length
125728 128545: contig of 2817 bp in length
128545 128645: gap of unknown length
128645 130563: contig of 1918 bp in length
130563 130663: gap of unknown length

```

* 130664 132945: contig of 2282 bp in length
* 132946 133045: gap of unknown length
* 133046 134556: contig of 1511 bp in length
* 134557 134656: gap of unknown length
* 134657 136478: contig of 1822 bp in length
* 136479 136578: gap of unknown length
* 136579 138159: contig of 1581 bp in length
* 138160 138259: gap of unknown length
* 138260 140456: contig of 2197 bp in length
* 140457 140556: gap of unknown length
* 140557 142544: contig of 1988 bp in length
* 142545 142644: gap of unknown length
* 142645 143784: contig of 1140 bp in length
* 143785 143884: gap of unknown length
* 143885 145304: contig of 1420 bp in length
* 145305 145404: gap of unknown length
* 145405 147391: contig of 1987 bp in length
* 147392 147491: gap of unknown length
* 147492 148908: contig of 1417 bp in length
* 148909 149008: gap of unknown length
* 149009 150147: contig of 1139 bp in length
* 150148 150247: gap of unknown length
* 150248 151752: contig of 1505 bp in length
* 151753 151852: gap of unknown length
* 151853 153441: contig of 1589 bp in length
* 153442 153541: gap of unknown length
* 153542 154953: contig of 1412 bp in length
* 154954 155053: gap of unknown length
* 155054 156244: contig of 1191 bp in length
* 156245 156344: gap of unknown length
* 156345 158095: contig of 1751 bp in length
* 158096 158195: gap of unknown length
* 158196 159309: contig of 1114 bp in length
* 159310 159409: gap of unknown length
* 159410 160682: contig of 1273 bp in length
* 160683 162602: gap of unknown length
* 162602 162702: contig of 1820 bp in length
* 162603 162702: gap of unknown length

```

Query Match 65.8%; Score 20.4; DB 2; Length 179030;
 Best Local Similarity 80.0%; Pred. No. 4e+02;
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

2 gccgcgggggtaataactacgccccgcgcgc 31
 12057 GCCCGGTGGGTACAGTACGCCGCGGCT 12086

RESULT 10
 AF005765/c 901 bp DNA linear BCT 15-JAN-2002
 LOCUS
 DEFINITION Uncultured archaeon Arc.171 16S ribosomal RNA gene, partial
 sequence.
 ACCESSION AF005765
 VERSION AF005765.2 GI:18152804
 KEYWORDS
 SOURCE uncultured archaeon Arc.171.
 ORGANISM uncultured archaeon Arc.171.
 Archaea; environmental samples.
 REFERENCE 1 (bases 1 to 901)
 AUTHORS Chandler,D.P., Brockman,F.J., Bailey,T.J. and Fredrickson,J.K.
 TITLE Phylogenetic Diversity of Archaea and Bacteria in a Deep Subsurface
 Paleosol
 JOURNAL Microb. Ecol. 36 (1), 37-50 (1998)
 PUBMED 9622563
 REFERENCE 2 (bases 1 to 901)
 AUTHORS Chandler,D.P.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAY-1997) Environmental Microbiology, Battelle
 Pacific Northwest National Lab, 900 Battelle Blvd., Mail Stop
 K4-06, Richland, WA 99352, USA
 COMMENT On Jan 15, 2002 this sequence version replaced gi:2443507.
 FEATURES Location/Qualifiers

source 1..901
 /organism="uncultured archaeon Arc.171"
 /db_xref="taxon:62249"
 rRNA
 complement(<1..>901)
 /product="16S ribosomal RNA"
 BASE COUNT 169 a 296 c 215 g 221 t
 ORIGIN

Query Match 65.2%; Score 20.2; DB 1; Length 901;
 Best Local Similarity 88.0%; Pred. No. 1.6e+03;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 agccgcggggtaataactacgcccc 25
 Db 420 AGCCGCGCGGTATATACGACCC 396

RESULT 11
 AF004344 1122 bp DNA linear BCT 25-MAR-1998
 LOCUS
 DEFINITION Uncultured crenarchaeote 16S ribosomal RNA gene, partial
 sequence.
 ACCESSION AF004344
 VERSION AF004344.1 GI:2984608
 KEYWORDS
 SOURCE uncultured crenarchaeote.
 ORGANISM uncultured crenarchaeote.
 Archaea; Crenarchaeota; environmental samples.

REFERENCE 1 (bases 1 to 1122)
 AUTHORS Vetrlian,C., Reysenbach,A.L. and Dore,J.
 TITLE Recovery and phylogenetic analysis of archaeal RNA sequences from
 continental shelf sediments
 JOURNAL FEMS Microbiol. Lett. 161 (1), 83-88 (1998)
 MEDLINE 98222521

REFERENCE 2 (bases 1 to 1122)
 AUTHORS Vetrlian,C., Reysenbach,A.-L. and Dore,J.
 TITLE Direct Submission
 JOURNAL Submitted (16-MAY-1997) Center of Marine Biotechnology, Maryland
 Biotechnology Institute, 701 E. Pratt St., Suite 236, Baltimore, MD
 2102, USA

FEATURES
 source 1..1122
 Location/Qualifiers

RNA
 BASE COUNT 273 a 273 c 367 g 209 t
 ORIGIN
 1..1122
 /organism="uncultured crenarchaeote"
 /db_xref="taxon:29281"
 /clone="BBA4"
 /note="Isolated from Buzzards Bay sediments, NW Atlantic
 ocean, depth 12 m
 type: Clone BBA4"
 <1..>1122
 /product="16S ribosomal RNA"

Query Match 65.2%; Score 20.2; DB 1; Length 1122;
 Best Local Similarity 88.0%; Pred. No. 1.5e+03;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 agccgcggggtaataactacgcccc 25
 Db 134 AGCCGCGCGGTATATACGACCC 158

RESULT 12
 AF361212 1411 bp DNA linear BCT 06-SEP-2001
 LOCUS
 DEFINITION Uncultured crenarchaeote clone SUBT-13 16S ribosomal RNA gene,
 partial sequence.
 ACCESSION AF361212
 VERSION AF361212.1 GI:14028778
 KEYWORDS

```

ORIGIN
Query Match          64.5%; Score 20; DB 1; Length 804;
Best Local Similarity 82.1%; Pred. No. 2e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 agccgcgggggtaatactagaccgccgcgcg 28
    ||||| ||||| || |||||
Db 431 AGCCGCCGCGGTATACCCAGCACCCAG 458

RESULT 14
ARI172771 31 bp DNA linear PAT 17-DEC-2001
LOCUS ARI172771
DEFINITION Sequence 2 from patent US 6303345.
ACCESSION ARI172771
VERSION ARI172771.1 GI:117912262
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Rohde,W., Becker,D., Randles,J.W., Hehn,A. and Salamini,F.
TITLE Use of a virus DNA as promoter.
JOURNAL Patent: US 6303345-A 2 16-OCT-2001;
FEATURES
source 1..31
location/Qualifiers
BASE COUNT 5 a 11 c 11 g 4 t
ORIGIN

Query Match          63.9%; Score 19.8; DB 6; Length 31;
Best Local Similarity 77.4%; Pred. No. 4.9e+03;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 agccgcgggggtaatactagaccgccgcgcgct 31
    ||||| ||||| | ||||| |||||
Db 31 AGCCGCCGCGGTATACCCAGCGGCT 1

RESULT 15
AF112330 519 bp DNA linear VRT 26-FEB-2001
LOCUS AF112330
DEFINITION Galaxias paucispombylus 16S ribosomal RNA gene, mitochondrial gene
for mitochondrial RNA, partial sequence.
ACCESSION AF112330
VERSION AF112330.1 GI:7688534
KEYWORDS
SOURCE Galaxias paucispombylus.
ORGANISM Mitochondrion Galaxias paucispombylus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Osmeriformes; Galaxiidae; Galaxias.
1 (bases 1 to 519)
Waters,J.M., Lopez,J.A. and Wallis,G.P.
Molecular phylogenetics and biogeography of galaxiid fishes
(Osteichthyes: Galaxiidae): dispersal, vicariance and the position
of Lepidogalaxias salamandroides
Syst. Biol. 49 (4), 777-795 (2000)
2 (bases 1 to 519)
Waters,J.M. and Wallis,G.P.
Direct Submission
Submitted (08-DEC-1998) Department of Zoology, University of Otago,
PO Box 56, Dunedin, Otago, New Zealand
location/Qualifiers
1..519
/organism="Galaxias paucispombylus"
/organellar="mitochondrion"
/db_xref="taxon:89560"
<1..>519
/product="16S ribosomal RNA"

```

Wed Jun 26 08:07:42 2002

us-09-462-955b-2.rge

Page 10

BASE COUNT 144 a 126 c 129 g 120 t
ORIGIN

Query Match 63.9%; Score 19.8; DB 5; Length 519;
Best Local Similarity 91.3%; Pred. No. 2.6e+03;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 ccgcgggggtataactagcccc 25
|||||||
Db 288 CCGCGGGGGAAGACTAGCCCC 310

Search completed: June 25, 2002, 15:39:46
Job time: 3607 sec